DOCUMENT RESUME

ED 188 891 SE 031 136

AUTHOR Tobin, Kenneth G.; Capie, William
TITLE The Test of Logical Thinking: Development and Applications.
PUB DATE Apr 80
NOTE 19p.; Paper presented at the annual meeting of the National Association for Research in Science Teaching (53rd, Boston, MA, April 11-13, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Measurement; *Cognitive Tests; *College Science; Developmental Stages; Higher Education; Science Education; Secondary Education; *Secondary School Science; *Test Construction; Tests
IDENTIFIERS *Piaget (Jean); Science Education Research

ABSTRACT

Traced are the developmental stages involved with the design of the Test of Logical Thinking (TOLT), a paper and pencil test constructed by the authors to measure formal reasoning ability. The final version of TOLT is described as consisting of two items for each of five reasoning modes: (1) controlling variables, (2) probabilistic reasoning, (3) combinatorial reasoning, (4) correlational reasoning, and (5) proportional reasoning. Further, a color videotape is utilized to present the context for each problem and to standardize administration procedures. Time is fixed as 40 minutes. Research is reported that supports TOLT as an effective means of identifying students of differing formal reasoning ability in classroom based research. Suggestions are made for design of further experiments as well as for application of the test to classroom instruction. (CS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kenneth G. Tobin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED188891

THE TEST OF LOGICAL THINKING:

DEVELOPMENT AND APPLICATIONS

Kenneth G. Tobin
William Capie

Department of Science Education
University of Georgia
Athens, Georgia 30602

A Paper Presented
at the Annual Meeting of the
National Association for Research in Science Teaching

Boston, MA

April, 1980

SE 031 136

THE TEST OF LOGICAL THINKING: DEVELOPMENT AND APPLICATIONS

INTRODUCTION

Much has been written on the importance of cognitive development as a variable that mediates science learning. Research has indicated that formal thought is required to learn many of the concepts taught in middle school, high school and college science courses (e.g., Goodstein and Howe, 1978; Cantu and Herron, 1978). Evidence also suggests that the majority of students in middle and high school grades (cf Chiappetta, 1976) and a significant proportion of college students (cf Renner and Lawson, 1973) are unable to utilize formal operations in problem solving.

The concensus of findings from research on cognitive development (e.g., Lawson, 1979; Lawson, Karplus and Adi, 1979; Lovell, 1961) is that science curricula need to be changed so that the cognitive development of learners becomes a focus; teachers need to match instruction to the cognitive level of learners; and research is needed to investigate the nature of learning for students at different levels of cognitive development. To attend to these concerns, researchers and teachers require a means of obtaining reliable and valid measures of cognitive development in group settings. In most cases the interview techniques that have been used in previous research are not an option because of the considerable expertise required by the interviewer and the time required to assess subjects on an individaul basis.

Because of the advantages of simultaneously assessing large numbers of subjects, a number of attempts have been made to construct paper and pencil measures of cognitive development (e.g., Staver and Gabel, 1979; Lawson, 1978: Tisher and Dale, 1975; Burney, 1974). The test described by

Lawson contained items that were selected to provide a measure of formal reasoning. Lawson identified five reasoning modes that are involved in "the search for and evaluation of evidence to support or reject hypothetical causal propositions". The five reasoning modes were identifying and controlling variables, combinatorial reasoning, probabilistic reasoning, correlational reasoning and proportional reasoning. A feature of the items used by Lawson was the requirement that evidence be used to justify the solution provided for each problem. Subjects were required to select an answer from a set of alternatives and provide written justification for the selected solution. This general approach and a selection of the items described by Lawson (1978) and Lawson et al (1979) were used as a basis for the development of the Test of Logical Thinking (TOLT) (Tobin and Capie, 1980a).

Treagust (1979) has argued that choice together with a reason was not indicative of formal thinking. He stated that "the only valid way to assess reasoning is by questioning the response". This represents an extreme viewpoint that would deny the possibility of assessing formal operations with pencil and paper measures. By designing a test that required a choice together with supporting reason, on multiple occasions for each mode of reasoning, the total test score should provide a measure of formal reasoning. The probability of a subject solving and justifying a significant number of items without using formal reasoning was considered to be small.

PURPOSE

The purpose of this paper is to describe the characteristics of TOLT and to suggest applications of the test for researchers and teachers.

DEVELOPMENT OF TOLT

Priority was directed to the development of a pencil and paper test of formal reasoning ability with sound psychometric properties. Items that had been

used in prior research (Lawson, 1978; Lawson et al, 1979) were used as a basis for developing an initial version of TOLT. The adoption of this procedure assured that TOLT would contain items that had been previously reported as valid measures of formal reasoning ability. Two items were selected to measure each of five modes of formal reasoning: controlling variables, proportional reasoning, probabilistic reasoning, correlational reasoning and combinatorial reasoning. The test incorporated demonstrations to provide a context for the items. Students selected a correct response from a number of alternatives and provided written justification for their selection.

The ten items were administered to a sample of high school students to obtain data on the types of justification provided for each selected response. Although the test reliability was reasonably high (α = .74) several factors reduced validity. For example, many students were unable to formulate a clear written justification for selecting a particular response. Scoring was also difficult; in fact, rescoring suggested many inconsistencies in the original scoring.

Because of these problems, the test was modified so that multiple justifications as well as multiple solutions were provided for each item. Written justifications provided by students in the initial trial were used as a basis for designing a set of plausible justifications. By incorporating responses previously volunteered as justifications, each justification was expected to appeal as a satisfactory explanation for a specific class of subjects. A correct solution to an item required that the correct response and the best justification for that response were selected.

The final version of TOLT consisted of two items for each of the five reasoning modes. A color video-tape was utilized to present the context for each problem and to standardize administration procedures. The time

required for the test to be administered was 40 minutes. A sample item is provided in Figure 1 as an example of the format used in the test.

(Include Figure 1 About Here)

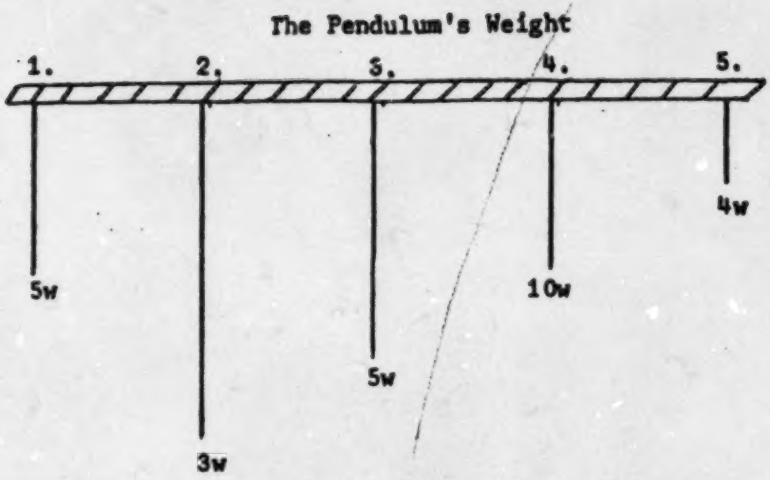
The ten TOLT items were selected and modified to provide a measure of formal reasoning ability. Data reported by Tobin and Capie (1980a) suggested that the ten items were measuring a common underlying dimension. Evidence of this was a high internal consistency reliability (α = .85) and a comparatively strong one-factor solution obtained from factor analysis of the achievement data on each of the reasoning modes. The moderately high correlations of each of the five reasoning modes with the factor provided an indication that the TOLT items were each related to the underlying dimension.

The major purpose in constructing TOLT was to obtain continuous measures of formal reasoning ability. Although raw scores have been used in research applications of TOLT, two alternatives should be investigated. The use of factor score coefficients from the factor analysis reported by Tobin and Capie enable TOLT performance to be expressed in terms of factor scores. Since the factor scores take account of variations in the strength of the relationships of the TOLT items with the underlying dimension, they may provide a better estimate of formal reasoning ability. The factor score coefficients are provided in Table 1.

(Include Table 1 About Here)

Latent trait models (cf Wright, 1977) provide an alternative to classical procedures for analyzing tests such as TOLT. Using a one parameter latent trait model, item difficulty indices and person ability measures were obtained. The person ability measures which are reported in Table 2 represent measures of performance on the "latent trait" which underlies TOLT

Item 7



Suppose you wanted to do an experiment to find out if changing the weight on the end of the string changed the amount of time the pendulum takes to swing back and forth. Which pendulums would you use for the experiment?

a. 1 and 4

b. 2 and 4

c. 1 and 3

d. 2 and 5

e. all

Reason

1. The heaviest weight should be compared to the lightest weight.

2. All pendulums need to be tested against one another.

3. As the number of washers is increased the pendulum should be shortened.

4. The number of washers should be different but the pendulums should be the same length.

5. The number of washers should be the same but the pendulums should be different lengths.

Figure 1. A sample item from TOLT

TABLE 1

Factor Score Coefficients for each TOLT Item

Item Number	Factor Score Coefficient
1	.183
2	.197
3	.182
4	.157
5	.133
6	.093
7	.093
8	.095
9	.141
10	.122

performance. Research is needed to determine whether raw scores, factor scores or latent trait scores provide the most valid measure of formal reasoning ability.

(Include Table 2 About Here)

Data reported by Tobin and Capie (1980a) allow subjects to be categorized as concrete, transitional or formal on the basis of TOLT performance. Although the categories obtained could be used as a basis for grouping subjects for analyses such as ANOVA, the authors see no advantage in adopting this procedure. The purpose in providing the data was to allow comparison of results obtained using TOLT with results obtained from studies in which interviews or other pencil and paper measures were used to obtain cognitive development data.

RESEARCH APPLICATIONS OF TOLT

TOLT has application in classroom based research. Tobin and Capie (1980b) reported formal reasoning ability, measured with TOLT, to be significantly related to summative achievement and retention of integrated science processes with middle school learners. At college level, Yeany, Helseth and Barstow (1980) found TOLT performance to be significantly related to genetics achievement. In the same investigation an interaction of formal reasoning ability with locus of control was reported on genetics achievement.

Capie and Tobin (in press) argued the importance of student engagement rates as a variable influencing achievement. It is apparent from cognitive development research cited earlier that high engagement rates alone cannot assure high achievement. Formal reasoning ability is likely to mediate achievement. As a consequence, formal reasoning ability should be used as a covariate in learning studies.

TABLE 2

Latent Trait Ability Estimates
for Total TOLT Scores

Total Score on TOLT	Latent Trait Ability Estimate
10	not estimable
9	2.35
8	1.50
7	0.92
6	0.44
5	0.00
4	-0.45
3	-0.92
2	-1.50
1	-2.34
0	not estimable

TOLT can also be used in aptitude-treatment-interaction studies. Tobin and Capie (1980c) reported a significant interaction of teacher wait-time with student formal reasoning ability on generalizing behavior of students in science classes. Students with higher levels of formal reasoning ability were associated with higher proportions of generalizing behavior in classes where a long wait-time was used.

CLASSROOM APPLICATIONS OF TOLT

Researchers have argued that instruction is most effective when it is matched to the developmental level of learners (Lovell, 1961; Lawson, 1979). However, such diagnosis represents an impossible challenge unless teachers can reliably measure developmental level of learners. TOLT represents a convenient means of identifying students of differing formal reasoning ability. The data in Table 3, which were collected in an urban high school, highlight potential problems for science teachers. If the TOLT data had been available, they could have been used as a basis for forming groups with less variation in formal reasoning ability. A considerable range in TOLT performance characterized the distribution in each class. The teacher of these chemistry classes was able to take immediate advantage of the data by providing individualized instruction for students with low TOLT scores. Similarly, extension problems were given to students with high TOLT scores, particularly in chemistry II. Although the same syllabus was used in all classes the needs of most students in Chemistry III are likely to differ from those of most students in Chemistry II. A different approach to teaching chemistry may be required in each classroom. The ability range could be considerably reduced by interchanging students between the three classes to obtain more homogeneous groups.

(Include Table 3 About Here)

TABLE 3

Distribution of Formal Reasoning
Ability in Three Grade 11
Chemistry Classes

Score on TOLT	Chemistry I	Chemistry II	Chemistry III
10	4	1	8
9	3	1	2
8	8	0	6
7	3	1	3
6	1	0	1
5	3	3	3
4	3	0	2
3	4	2	0
2	3	1	0
1	1	5	1
0	2	7	0

At college level TOLT has been successfully used to match groups for research and instructional purposes. In these instances the objective was to obtain heterogeneous groups with similar distributions of formal reasoning ability.

APPLICATIONS OF THE METHODOLOGY USED TO DEVELOP TOLT

The methodology used to develop the ten TOLT items is applicable to other situations. Parallel forms of the test have already been developed for use in classroom based research. Another application is to develop diagnostic tests to separately investigate each reasoning mode. Figures 2-4 contain examples of items used to investigate different aspects of proportional reasoning.

(Include Figures 2, 3 and 4 About Here)

Another extension of the methodology is to utilize interactive computer terminals. This mode of assessment offers the flexibility of being able to match a question to be posed to the response and/or justification provided for a previous item. Although computers are not an alternative where large group assessment is a concern they do provide an opportunity for valid and reliable measures of formal reasoning to be obtained when smaller samples are involved or computer terminals are available in sufficient numbers.

CONCLUSION

TOLT provides a convenient means of obtaining valid and reliable measures of formal reasoning ability for researchers and teachers, particularly when data are required for groups of subjects. TOLT has many potential applications for research. For example, TOLT data can be used as a basis for stratifying subjects prior to random assignment to treatment groups; as a moderating variable in studies of teaching and learning; and as a dependent measure in studies that are concerned with variables that influence formal reasoning

Item 1 The Seesaw

A boy sits 6 feet from the center of a seesaw. The seesaw is balanced when his father sits 4 feet from the center on the other side of the seesaw.

The father moves so that he is 2 feet from the center. Where must the boy sit to balance the seesaw?

a. 1 foot from the center
b. 2 feet from the center
c. 3 feet from the center
d. 4 feet from the center
e. 5 feet from the center

Reason

1. The boy should be closer to the center.
2. The boy and his father should be equal distances from the center.
3. Since the father halved his distance from the center, the boy should do the same.
4. Since the father moved 2 feet closer to the center the boy should do the same.
5. The boy should always be 2 feet further from the center than his father.

Figure 2. Proportional reasoning example 1

Item 4

Orange Juice #1

Four large oranges are squeezed to make six glasses of juice. How much juice can be made from six oranges?

a. 7 glasses
b. 8 glasses
c. 9 glasses
d. 10 glasses
e. other

Reason

1. The number of glasses compared to the number of oranges will always be in the ratio 3 to 2.
2. With more oranges, the difference will be less.
3. The difference in the numbers will always be two.
4. With four oranges the difference was 2. With six oranges the difference would be two more.
5. There is no way of predicting.

Figure 3. Proportional reasoning example 2

Item 5 Orange Juice #2

How many oranges are needed to make 13 glasses of juice?

a. 6 1/2 oranges
b. 8 2/3 oranges
c. 9 oranges
d. 11 oranges
e. other

Reasons

1. The number of oranges compared to the number of glasses will always be in the ratio 2 to 3.
2. If there are seven more glasses, then five more oranges are needed.
3. The difference in the numbers will always be two.
4. The number of oranges will be half the number of glasses.
5. There is no way of predicting the number of oranges.

Figure 4. Proportional reasoning example 3

variable in studies of teaching and learning; and, as a dependent measure in studies that are concerned with variables that influence formal reasoning abilities of learners.

The methodology used to develop TOLT items is generally applicable to test development. Of particular interest are applications to produce parallel forms of TOLT; to produce items that further probe the nature of formal reasoning; and, to develop an interactive computer test for assessing formal reasoning ability.

REFERENCES

Burney, G. M. The construction and validation of an objective formal reasoning instrument. Unpublished doctoral dissertation, University of Northern Colorado, 1974.

Cantu, L. & Herron, D. Concrete and formal Piagetian stages and science concept attainment. Journal of Research in Science Teaching, 1978, 15, 135-143.

Capie, W. & Tobin, K. G. Pupil engagement in learning tasks: A fertile area for research in science teaching. Journal of Research in Science Teaching, (in press).

Chiappetta, E. A review of Piagetian studies relevant to science instruction at the secondary and college level. Science Education, 1976, 60, 253-261.

Goodstein, M. & Howe, A. The use of concrete methods in secondary chemistry instruction. Journal of Research in Science Teaching, 1978, 15, 361-366.

Lawson, A. E. The development and validation of a classroom test of formal reasoning. Journal of Research in Science Teaching, 1978, 15, 11-24.

Lawson, A. E. Combining variables, controlling variables and proportions: Is there a psychological link? Science Education, 1979, 63, 67-72.

Lawson, A. E., Adi, H. & Karplus, R. Development of correlational reasoning in secondary schools: Do biology courses make a difference? The American Biology Teacher, 1979, 41, 420-425.

Lovell, K. A follow-up study of Inhelder and Piaget's the growth of logical thinking. British Journal of Psychology, 1961, 52, 143-153.

Renner, J. W. & Lawson, A. E. Promoting intellectual development through science teaching. Physics Teacher, 1973, 11, 273-276.

Staver, J. R. & Gable, D. L. The development and construct validation of a group-administered test of formal thought. Journal of Research in Science Teaching, 1979, 16, 535-544.

Tisher, R. P. & Dale, L. G. Understanding in science test. Victoria: Australian Council for Educational Research, 1975.

Tobin, K. G. & Capie, W. The development and validation of a group test of logical thinking. Paper presented at the American Educational Research Association Meeting, Boston, April, 1980a.

Tobin, K. G. & Capie, W. The effects of student engagement rate, formal reasoning ability and locus of control on middle school science achievement. Paper presented at the Third Annual Conference of the Eastern Educational Research Association, Norfolk, Virginia, March, 1980b.

REFERENCES (cont.)

Tobin, K. G. & Capie, W. Effects of teacher wait-time in middle school science. Paper to be presented at the Annual Meeting of the Americal Psychological Association, Montreal, Canada, September, 1980c.

Treagust, D. F. Comments on "The acquisition of propositional logic and formal operational schemata during the secondary school years". Journal of Research in Science Teaching, 1979, 16, 363-367.

Wright, B. D. Solving measurement problems with the Rasch model. Journal of Educational Measurement, 1977, 14, 97-116.

Yeany, R. H., Helseth, E. A. & Barstow, W. E. Interactive instructional video-tapes, scholastic aptitude, cognitive development and locus of control as variables influencing science achievement. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Boston, MA, April, 1980.